Lokalisierung der Leber mittels einer Diskriminativen Generalisierten Hough Transformation

Heike Ruppertshofen[1,2],Cristian Lorenz[3],Sarah Schmidt[4,2], Peter Beyerlein[4], Zein Salah[2], Georg Rose[2], Hauke Schramm[1]

[1] Fachhochschule Kiel, Institut für Angewandte Informatik, Kiel, Germany
[2] Otto-von-Guericke Universität, Institut für Elektronik, Signalverarbeitung und Kommunikationstechnik, Magdeburg, Germany
[3] Philips Research Europe – Hamburg, Department Digital Imaging, Hamburg, Germany
[4] Technische Hochschule Wildau, Fachbereich Ingenieurwesen, Wildau, Germany

Kontakt: heike.ruppertshofen@fh-kiel.de

Abstract:

In diesem Beitrag soll ein für 2D Bilder bereits erfolgreich eingesetztes Verfahren für die automatische Objektlokalisierung auf 3D Problematiken erweitert und die Einsetzbarkeit des Verfahrens für 3D Daten gezeigt werden. Das Verfahren der Objektlokalisierung basiert dabei auf der Generalisierten Hough Transformation (GHT), bei der das dort verwendete Formmodell mit individuellen Punktgewichten ausgestattet wird, die im Abstimmverfahren der GHT zum Einsatz kommen und so die Genauigkeit des Verfahrens verbessern. Diese Punktgewichte werden mittels eines diskriminativen Lernverfahrens bezüglich eines minimalen Lokalisierungsfehlers in der GHT optimiert. Die trainierten Punktgewichte erhöhen dabei nicht nur die Genauigkeit der GHT, sondern bewerten auch die Wichtigkeit eines Punktes, so dass unwichtige Punkte aus dem Modell entfernt werden können. Kleinere Modelle führen zu kürzen Laufzeiten, was den Einsatz des Verfahrens auch für 3D Aufgabenstellungen praktikabel macht. Die Stärke des Verfahrens soll hier am Beispiel der Lokalisierung der Leber, die meist eine recht starke Variabilität aufweist, gezeigt werden.

Schlüsselworte: Objektlokalisierung, Generalisierte Hough Transformation, Diskriminatives Training, Maschinelles Lernen

1 Problem

Die Aufgabe der Objektlokalisierung kommt bei vielen Anwendungen in der computergestützten Medizin - im Rahmen der Diagnostik, der Radiologie und der Chirurgie - zum Tragen. Viele der Verfahren, insbesondere in der Operations- und Bestrahlungsplanung, stützen sich dabei auf eine vorherige Segmentierung von Organen oder Knochen, um voll automatisiert Volumen, Längen oder Winkel zu berechnen und die Planung einer Operation oder Bestrahlung am Computer zu ermöglichen.
Die meisten dieser automatischen Segmentierungsverfahren benötigen zunächst die ungefähre Position des Zielorgans, um ein Segmentierungsmodell initial zu positionieren. In vielen Fällen wird diese Position manuell bestimmt oder es kommen speziell auf das gestellte Problem zugeschnittene Verfahren zum Einsatz [4]. Die Anpassung dieser Verfahren auf neue Problemstellungen ist zeitaufwändig und häufig nicht machbar.
In diesem Beitrag wollen wir daher ein generelles und automatisches Verfahren zur Objektlokalisierung vorstellen, das problemlos auf weitere Objekte angewendet werden kann. Das Verfahren besteht aus zwei Modulen: Zum einen der Generalisierten Hough Transformation (GHT) [2] zur Objektlokalisierung, die hier insofern erweitert wurde, dass ein gewichtetes Modell verwendet wird, und zum anderen ein diskriminatives Lernverfahren (DMC) [3], mit dessen Hilfe die individuellen Gewichte für die einzelnen Modellpunkte ermittelt werden. Durch das Training wird erreicht, dass Bereiche des Modells, die besonders charakteristisch für das Zielobjekt sind, stärker gewichtet werden und so auch einen stärkeren Einfluss in der GHT haben.
Auf 2D Bildern kommt das Verfahren bereits erfolgreich zum Einsatz [1]; In diesem Beitrag soll nun die Einsetzbarkeit des Verfahrens auf 3D Bildern gezeigt werden. Als Anwendungsbeispiel wird die Lokalisierung der Leber in 3D CT-Aufnahmen verwendet. Um die Leistung der neuen Methode im direkten Vergleich zur Standard-GHT darstellen zu können, wird die Lokalisierung mit einem mittleren Formmodell, welches die durchschnittliche Form der Leber repräsentiert, durchgeführt. Die Erstellung eines in der GHT verwendbaren Modells für das Zielobjekt direkt aus den Daten, wie es in [1] für den 2D-Fall vorgestellt wurde, ist aber ebenfalls möglich.

2 Methoden und Material

Wie bereits in der Einleitung erwähnt wurde, besteht das Verfahren aus zwei Modulen, auf die im Folgenden näher eingegangen werden soll.
Für die Objektlokalisierung wird die GHT [2] eingesetzt. Das Verfahren eignet sich gut für medizinische Bildverarbeitung, da es robust gegenüber Bildrauschen und Überdeckungen oder fehlenden Objekteilen ist. Nachteil des Verfahrens ist die hohe Laufzeit, die vermutlich dafür verantwortlich war, dass die GHT auf 3D Problematiken kaum Anwendung fand, und die hier durch kleinere Modelle und stark eingeschränkte Transformationsparameter reduziert werden soll.
Die GHT führt die Lokalisierung mittels eines Abstimmverfahrens durch, bei dem das Bild mit Hilfe des Modells in den sogenannten Hough-Raum transformiert wird. Zu diesem Zweck wird ein Kantenbild des Originalbildes erstellt und für jeden Kantenpunkt e_j die Modellpunkte m_k mit ähnlicher Gradientenrichtung ermittelt. Über den Zusammenhang $c_i = e_j - m_k$ wird die korrespondierende Hough-Zelle c_i ermittelt und ihr Wert erhöht. Nach Beendigung des Abstimmverfahrens wird die Zelle mit dem höchsten Wert gesucht und als wahrscheinlichste Objektposition ausgegeben. Das Verfahren kann weiterhin verwendet werden, um skalierte und rotierte Vorkommen des Zielobjekts zu lokalisieren. Dafür muss das Modell transformiert und das Verfahren wiederholt werden, was zu einem höher dimensionalen Hough-Raum führt. Aus Laufzeitgründen werden diese Möglichkeiten hier jedoch nicht in Betracht gezogen. Stattdessen soll das Modell lernen, welche Bereiche für die Lokalisierung am verlässlichsten und wichtigsten sind, so dass die Bestimmung der Position des Zielobjektes im Bild auch ohne Schätzung weiterer Parameter erfolgreich ist.
Die Erweiterung der GHT von 2D auf 3D ist unkompliziert und benötigt lediglich eine Erweiterung der Datenstrukturen und eine speichersparende Verwaltung des Hough-Raumes und der für das folgende Gewichttraining benötigten Informationen in Heap- und Hashstrukturen.
In der Standard-GHT haben alle Modellpunkte einen gleich starken Einfluss auf die Lokalisierung. Bei jedem Objekt gibt es jedoch Gebiete, die charakteristisch sind und somit in allen Bildern ähnlich aussehen, und wiederum andere die recht stark variieren. Dieses soll hier im Modell berücksichtigt werden, indem für jeden Modellpunkt ein individuelles Gewicht bestimmt wird, welches im Abstimmprozess der GHT verwendet werden soll. Dabei sollen Punkte, die das Zielobjekt robust beschreiben, höhere Gewichte bekommen und Punkte, die nur bei sehr wenigen Bildern passen, ein niedriges Gewicht erhalten, so dass sie später über ihr Gewicht identifiziert und aus dem Modell entfernt werden können. Des Weiteren können auch negative Gewichte vergeben werden, um das Modell von verwechselbaren Strukturen abzustoßen und so eine falsch-positive Lokalisierung zu vermeiden.
Für die Ermittlung dieser Modellpunktgewichte wird ein diskriminatives Lernverfahren [3] herangezogen, das erstmalig in der Spracherkennung eingesetzt wurde. Bei diesem Lernverfahren wird die Dimensionalität der Bilder nicht berücksichtigt, so dass es direkt für die drei dimensionalen Bilder anwendbar ist. In dem Verfahren wird jeder Modellpunkt mit seinem Beitrag zum Hough-Raum als individuelle Wissensquelle aufgefasst und diese Wissensquellen über einen log-linearen Ansatz kombiniert. Die über den log-linearen Ansatz eingeführten Gewichte für die einzelnen Wissensquellen sollen hier als Modellpunktgewichte verwendet werden. Um diese bestimmen zu können, wird eine Fehlerfunktion definiert, die den Fehler in Abhängigkeit der Gewichte über alle Trainingsbilder akkumuliert und so eine nahezu optimale Gewichtung der Wissensquellen ermittelt, bei der der Fehler über alle Bilder am kleinsten ist. Für eine detailierte Herleitung und Erklärung der Bestimmung der Modellpunktgewichte verweisen wir den Leser auf [1].
Nach Ermittlung der Gewichte kann die Modellgröße reduziert werden, indem Modellpunkte mit einem kleinen absoluten Gewicht aus dem Modell entfernt werden. Gleichzeitig wird das Modell diskriminativer für das Zielobjekt, da robuste Modellpunkte einen stärkeren Einfluss bekommen und verwechselbare Modellbereiche entfernt werden.
Die Methode wird auf 38 CT-Aufnahmen des Rumpfes getestet mit der Aufgabe, die Leber zu lokalisieren. Die verwendeten Daten stammen aus einem Leber-Segmentierungs-Challenge [5] sowie einer öffentlich verfügbaren Datenbank [6], bei der auf etwa 75% der Bilder Lebertumore sichtbar sind. Die Bilder haben unterschiedliche Qualität und Auflösung, die zwischen $0.56 \times 0.56 \times 1$ mm und $0.87 \times 0.87 \times 4$ mm schwankt. Ein koronaler Schnitt durch einen der Datensätze ist in Abb. 1, links zu sehen. Aus Laufzeitgründen wurden die Bilder für das Verfahren zwei Mal heruntergesampelt; wird eine höhere Genauigkeit benötigt, kann dieser Schritt jedoch übersprungen oder ein mehrstufiger Ansatz verfolgt werden.
Von den gegebenen Bildern werden 8 im Training der Modellpunktgewichte eingesetzt und 30 zu Testzwecken zurück behalten, um zu evaluieren, wie gut die Lokalisierungseigenschaften des Modells auf unbekannten Daten sind. Als Modell wird in diesem Beitrag ein mittleres Punktmodell der Leber mit 5120 Modellpunkten verwendet (s. Abb. 2, links). Als Referenzpunkt, bzw. Zielpunkt für die Leber wird ihr Schwerpunkt verwendet, der aus vorliegenden Segmentierungen der Daten berechnet wurde.

3 Ergebnisse

In Abb. 1 wird ein Beispielergebnis der Lokalisierung vorgestellt. Das linke Bild zeigt dabei einen koronalen Schnitt durch den Körper an der Stelle des gefundenen Punktes. Das rechte Bild stellt den dazugehörigen Schnitt durch

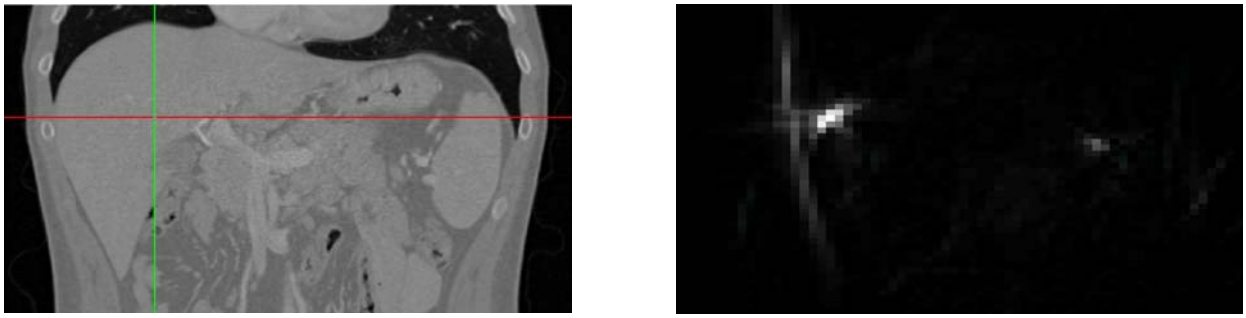

Abb. 1: Das linke Bild zeigt einen koronalen Schnitt durch eines der CT Bilder mit sichtbarer Leber. Der von der GHT gefundene Punkt wird durch die beiden Linien markiert. Das rechte Bild zeigt den dazugehörigen Hough-Raum der gleichen Schnittebene. Die Zellen mit den höchsten Werten sind deutlich zu erkennen.

den Hough-Raum dar mit einem klar sichtbaren Maximum. In diesem Beispiel wurde der Referenzpunkt in der Leber mit einer Genauigkeit von 0.7 cm lokalisiert. Abbildung 3 zeigt ein weiteres Beispiel, bei dem die Lokalisierung mit dem untrainierten und trainierten Modell verglichen wird.
Die weiteren Ergebnisse der Leberlokalisierung mit dem gewichteten und ungewichteten Modell sind in Tabelle 1 zu finden. Mit dem ungewichteten Modell wird die Leber in 18 von 38 Bildern gefunden mit einem mittlerem Fehler von 4.7 cm, bzw. 4.3 cm auf den Test-, bzw. Trainingsbildern. Durch die hohe Variabilität der Leber ist die Durchführung der Standard-GHT mit einem mittleren Modell in vielen Fällen nicht erfolgreich. Wird ein gewichtetes Modell verwendet erhöht sich die Genauigkeit der GHT auf 2.8 cm auf den unbekannten Testdaten und die Leber kann in allen Bildern lokalisiert werden. Durch eine Reduktion des Modells auf 500 Punkte werden unwichtige Punkte entfernt, wobei die Lokalisierungsgenauigkeit erhalten bleibt. Diese Ergebnisse sind sehr zufriedenstellend, insbesondere wenn man die grobe Auflösung des Hough-Raumes von durchschnittlich 0.5 cm und die vage Annotation der Zielpunkte, die sich nur am Schwerpunkt der Leber und nicht an klar definierten anatomischen Strukturen orientieren, in Betracht zieht.
Die gewichteten Modelle sind in Abb. 2 zu sehen. Dabei zeigt die linke Abbildung das komplette Modell, während die rechte Abbildung nur die 500 Punkte mit dem höchsten absoluten Gewicht darstellt. Aus dem rechten Bild wird ersichtlich, dass die obere Kante der Leber am robustesten für die Erkennung ist, da hier am wenigstens Variabilität auftritt. Hinzukommt, dass die Kante zwischen Leber und Lunge in allen Bildern gut zu erkennen ist. Die vordere Seite der Leber wurde mit negativen Gewichten versehen. Hierbei handelt es sich also um Punkte, die an einer anderen Stelle des Bildes besser passen und zu falschen Lokalisierungsergebnissen führen würden. Die laterale Seite der Leber erhielt nur niedrige Gewichte, da hier die Gefahr der Verwechslung mit der Körperaußenkontour besteht, wie in Abb. 3 für den Fall des untrainierten Modells gut zu erkennen ist.
Durch die Reduktion der Modellgröße verringert sich des Weiteren auch die Laufzeit des Verfahrens um 50%, wobei ca. 20s für Vorverarbeitungs- und Nachverarbeitungsschritte, wie z.B. das Ein- und Auslesen der Bilder, das Heruntersampeln und die Berechnung der Kanten, fallen. An dieser Stelle sollte dazu gesagt werden, dass das Verfahren bisher nicht laufzeitoptimiert ist. Insbesondere durch die gute Parallelisierbarkeit der verwendeten Methoden besteht hier noch ein hohes Potential für eine deutliche Reduktion der Laufzeit.

4 Diskussion

Der vorgestellte generelle Ansatz zur Objektlokalisierung zeigt auch auf 3D Daten zufriedenstellende Ergebnisse. Durch die individuelle Gewichtung der Modellpunkte konnte eine deutliche Erhöhung der Lokalisierungsgenauigkeit und eine

	Untrainiertes Modell	Trainiertes Modell	
Modellgröße	5120	5120	500
Fehler Trainingsbilder [cm]	4.3 ± 3.3 (10.1)	2.0 ± 1.2 (4.8)	2.1 ± 1.2 (4.8)
Fehler Testbilder [cm]	4.7 ± 2.8 (10.9)	2.8 ± 1.2 (6.1)	2.7 ± 1.2 (6.1)
Lokalisierungsrate [%]	47.4	100	100
Laufzeit [s]	49	49	24

Tabelle 1: Vergleich des ungewichteten und gewichteten Modells. Angegeben sind die Anzahl der Modellpunkte, der Lokalisierungsfehler der erfolgreichen Lokalisierungen in cm (Format: Mittelwert ± Standardabweichung (Maximum)), die Anzahl der erfolgreichen Lokalisierungen in % und die durchschnittliche Laufzeit pro Bild in Sekunden.

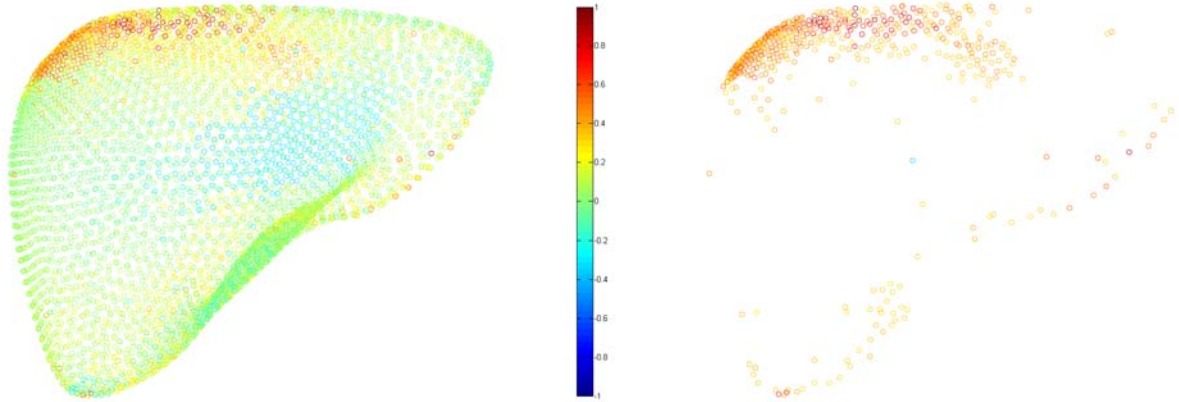

Abb. 2: Die Abbildung zeigt das mittlere Modell der Leber (links), wobei die Punkte entsprechend ihres Gewichts farbkodiert wurden (rot = 1, blau = -1). Das rechte Bild zeigt die 500 Punkte des Modells mit dem höchsten absoluten Gewicht.

Reduktion der falsch-positiven Lokalisierungen erreicht werden. Die Schätzung von Rotations- und Skalierungsparametern für die Lokalisierung des Zielobjektes wird dabei durch die Gewichtung des Modells redundant.
Ein großer Vorteil des Verfahrens besteht darin, dass es automatisch abläuft und ohne Anpassungen auf neue Aufgabenstellungen angewendet werden kann. Zwar wurde das Verfahren hier zu Vergleichszwecken unter Verwendung eines zuvor generierten, mittleren Modells vorgestellt, praktisch kann das Verfahren aber auch mit einer zufälligen Punktwolke arbeiten, da es selbständig die für das Zielobjekt relevanten Punkte identifiziert. Somit benötigt das Verfahren als Eingabe lediglich die Annotation des Zielpunktes in den Trainingsbildern, um ein diskriminatives Modell für die Lokalisierung zu erstellen.
Ein weiterer Vorteil der Methode liegt in der möglichen Reduktion der Modellgröße. Unwichtige Punkte, die keinen starken Beitrag zur Lokalisierung des Zielobjektes leisten, können anhand ihres niedrigen Gewichts identifiziert und aus dem Modell entfernt werden. Dieses führt zu einem klareren Ergebnis im Hough-Raum, da weniger störende Stimmen abgegeben werden, und einer deutlichen Reduktion der Laufzeit.
Das Verfahren wurde hier auf heruntergesampelten Bildern und einem relativ grob quantisierten Hough-Raum vorgestellt. Dieser hatte eine nahezu isotrope Auflösung von 0.5 cm im Mittel. Vor diesem Hintergrund ist die erreichte Lokalisierungsgenauigkeit von 2 cm im Mittel, was einem Fehler von 4 Voxeln entspricht, sehr zufriedenstellend. Sollte eine höhere Genauigkeit benötigt werden, kann jedoch auch mit voll aufgelösten Bildern gearbeitet oder ein mehrstufiger Ansatz verfolgt werden.
In diesem Beitrag sollte zunächst die Eignung des Verfahrens für 3D Bilder gezeigt werden. Um die Modelle, die in der GHT verwendet werden, weiter zu verbessern, soll als nächster Schritt die iterative Modellbildung wie in [1] zum Einsatz kommen, um die Stärken des Verfahrens voll nutzen zu können. Hierbei wird das Modell direkt aus den Daten generiert, um die vorliegende Variabilität in Rotation und Skalierung im Modell einfangen zu können. Gleichzeitig wird bei diesem Ansatz auch Information über verwechselbare Strukturen in das Modell gelernt, um die Gefahr falschpositiver Lokalisierungen weiter zu minimieren.

5 Referenzen

[1] H. Ruppertshofen, C. Lorenz, S. Strunk, P. Beyerlein, Z. Salah, G. Rose, H. Schramm, Fully Automatic Model Creation for Object Localization utilizing the Generalized Hough Transform, Proceedings of Bildverarbeitung für die Medizin, Springer, 2010

[2] D. H. Ballard, Generalizing the Hough Transform to Detect Arbitrary Shapes, Pattern Recognition 13(2), 1981

[3] P. Beyerlein, Discriminative Model Combination, Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing, 1998

[4] T. Heimann, B. van Ginneken, M. Styner et al., Comparison and Evaluation of Methods for Liver Segmentation from CT Datasets, IEEE Transactions on Medical Imaging 28(8), 2009

[5] B. van Ginneken, T. Heimann, M. Styner, 3D Segmentation in the Clinic: A Grand Challenge, Proceedings of the 3D Segmentation in the Clinic: A Grand Challenge Workshop of the 9th International Conference on Medical Image Computing and Computer-Assisted Intervention, 2007

[6] http://www.ircad.fr/softwares/3Dircadb/3Dircadb1/index.php?lng=en

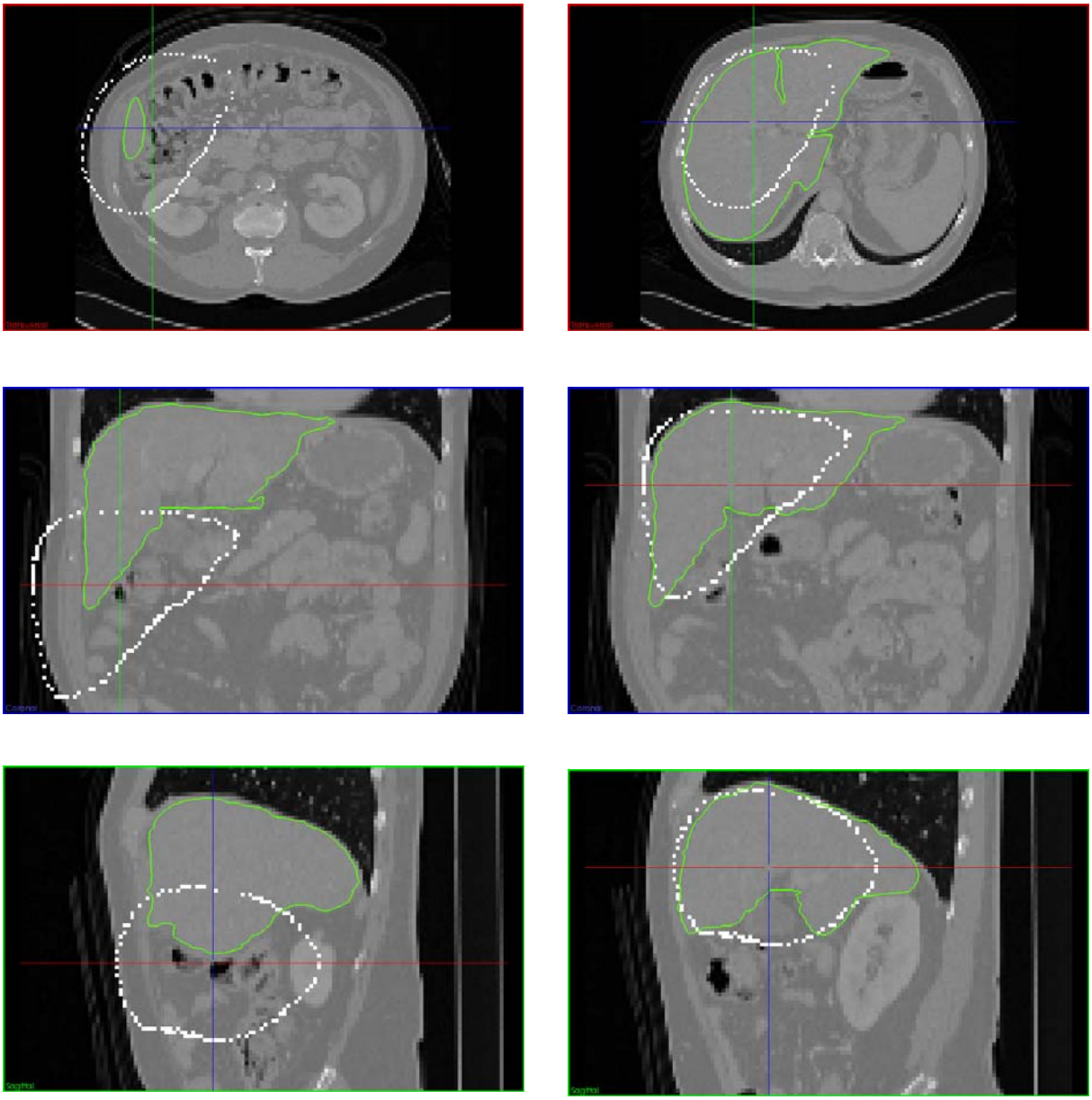

Abb. 3: Vergleich des Lokalisierungsergebnisses mit dem untrainierten Modell (links) und dem gewichteten (rechts) in axialer, koronaler und sagittaler Ansicht. Der von der GHT als bestes Ergebnis bestimmte Punkt wird durch das Fadenkreuz markiert. Zusätzlich sind das verwendete mittlere Modell (weiß) und die Segmentierung der Leber als Grundwahrheit (grün) eingezeichnet.